REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
21-08-2012	Conference Proceeding	-

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Capacity Bounds and Stochastic Resonance for Binary Input Binary Output Channels	W911NF-11-1-0144
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER 206022
6. AUTHORS Ira S. Moskowitz, Paul Cotae, Pedro N. Safier , Daniel L. Kang	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES	8. PERFORMING ORGANIZATION REPORT NUMBER
The University of the District of Columbia Graduate Studies University of the District of Columbia Washington, DC 20008 -0000	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211	10. SPONSOR/MONITOR'S ACRONYM(S) ARO 11. SPONSOR/MONITOR'S REPORT NUMBER(S) 58962-NS-REP.3

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
In this paper, we discuss the threshold based stochastic resonance behavior of binary input, binary output discrete memoryless channels. We follow the basic model of Chapeau-Blondeau where he showed how the addition of external Gaussian noise could enhance channel capacity, thus providing an optimum

15. SUBJECT TERMS
Bioinformatics and Biocomputing, Systems Securityand Trust, Computer and Network Security, Medical Applicationsand e-healthcare.

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
UU	UU	UU	UU		Paul Cotae
					19b. TELEPHONE NUMBER 202-274-6290

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

Report Title

Capacity Bounds and Stochastic Resonance for Binary Input Binary Output Channels

ABSTRACT

In this paper, we discuss the threshold based stochastic resonance behavior of binary input, binary output discrete memoryless channels. We follow the basic model of Chapeau-Blondeau where he showed how the addition of external Gaussian noise could enhance channel capacity, thus providing an optimum channel capacity depending on both the threshold level and noise power of the system. In order to easily approximate channel capacity behavior, we use capacity bounds and approximations.

Conference Name: Proceedings of the IEEE ComCompAp 2012, Hong Kong, China

Conference Date: January 11, 2012

Capacity Bounds and Stochastic Resonance for Binary Input Binary Output Channels

Ira S. Moskowitz*, Paul Cotae‡, Pedro N. Safier§, and Daniel L. Kang, student¶
*Naval Research Laboratory, Center for High Assurance Computer Systems - 5540, Washington, DC 20375
ira.moskowitz@nrl.navy.mil
‡Electrical and Computer Engineering Department, University of the District of Columbia, Washington, DC 20008
pcotae@udc.edu
§S&J Solutions LLC, 107 S. West St., PMB 509, Alexandria, VA 22314
pedro-safier@sj-solutions.com
¶Thomas Jefferson High School for Science and Technology, Alexandria, VA 22312
daniel.lee.kang@gmail.com

***Abstract*—In this paper, we discuss the threshold based stochastic resonance behavior of binary input, binary output discrete memoryless channels. We follow the basic model of Chapeau-Blondeau where he showed how the addition of external Gaussian noise could enhance channel capacity, thus providing an optimum channel capacity depending on both the threshold level and noise power of the system. In order to easily approximate channel capacity behavior, we use capacity bounds and approximations.**

***Keywords*—Bioinformatics and Biocomputing, Systems Security and Trust, Computer and Network Security, Medical Applications and e-healthcare.**

I. Introduction

The communication channels under study in this paper are binary input, binary output discrete memoryless channels [22], [24], [14], [15], denoted as (2,2) channels. That is, the channel is discrete, memoryless, the input random variable X has (non-trivial) values $\{\iota_1, \iota_2\}$, and the output random variable Y has (non-trivial) values $\{o_1, o_2\}$. Where (overloading notation), the event $X = \iota_k$ is the event that the symbol ι_k is input to the channel, and the event $Y = o_k$ is the event that the channel output is the symbol o_k.

We use the $(2,2)$ channel to study the phenomenon of stochastic resonance that arises from modeling the discrete time threshold neuron model [7]. A introduction to this is given in [9, Sec. 2]. We closely follow the mathematics of [3] in this paper.

For our purposes, stochastic resonance is described as a counterintuitive phenomenon in the presence of additive Gaussian noise. That is, under the proper conditions, for both small and large additive noise, associated performance metrics (e.g., SNR, mutual information, Fisher information, correlation, discrimination index) are small. However, for some intermediate noise levels, performance actually improves!

We focus on maximizing the mutual information, and finding the capacity, of (2,2) channels in threshold communication systems [1], [3]. We discuss that in these (2,2) channels there is a maximum of performance measured in terms of channel capacity. This maximum of channel capacity occurs at an optimal value of the added Gaussian noise correlated with the threshold level.

Our major contribution is to show how certain approximations to the closed form of channel capacity reflect physical intuition. This is of interest because even in the simple case of the $(2,2)$ channel, the non-linearity of the capacity expression makes closed from analysis intractable.

We also briefly show how to view stochastic resonance results in terms of the new area of *algebraic information theory* [14], [15], and provide geometric motivation for the closed form results. In particular, we show the significance of the determinant of the channel matrix in terms of the capacity approximations and bounds.

Note that an immediate application of the ideas in this paper is the counterintuitive paradigm that adding jamming noise [21] might actually increase communication throughput. Rather, if the jamming is not done correctly, then we might be strengthening the communication we are attempting to jam.

I-A. The (2,2) Channel

We summarize freely, without quotation marks, from [16] in this subsection. Since all symbols take the same time to go through the channel, all information theoretic measurements are in units of bits per channel usage (symbol). Furthermore, unless noted otherwise, all logarithmic quantities are base two logarithms.

The channel matrix represents the conditional probability relationships between the input and output symbols. That is, $a = P(o_1|\iota_1), 1-a = P(o_2|\iota_1), c = P(o_1|\iota_2)$, and $1-c = P(o_2|\iota_2)$. The input probabilities are represented by the random variable X, $P(X = \iota_i) = x_i, i = 1, 2$, which we simplify to $P(\iota_1) = x_1$ and $P(\iota_2) = x_2$. Similarly, we have the random variable Y such that $P(o_1) = y_1$ and $P(o_2) = y_2$. This is illustrated in Figure 1.

$$M = \begin{pmatrix} a & 1-a \\ c & 1-c \end{pmatrix}$$

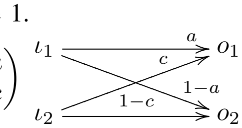

Fig. 1. $(2,2)$ Channel matrix and transition diagram. Note that $\det M = a - c$.

U.S. Government work not protected by U.S. copyright

We summarize the probabilities as follows

$$\vec{y} = \vec{x} \cdot \begin{pmatrix} a & \bar{a} \\ c & \bar{c} \end{pmatrix}.$$

Where $\bar{a} = 1 - a$, and $\bar{c} = 1 - c$. Letting $x = x_1$ and using the fact that $x_1 + x_2 = 1 = y_1 + y_2$, we have that $y_1 = (a - c)x + c$ and $y_2 = 1 - y_1$. To calculate the capacity we want to maximize the mutual information $I(X, Y)$

$$I = H(Y) - H(Y|X)$$

over all possible distributions of $(x_1, x_2) = (x, 1 - x)$.

Since, with a and c fixed, we can view I as a function of one variable x, the maximization problem reduces to maximizing the function $I : [0, 1] \to \mathbb{R}$ given by

$$I(x) = h\big(f(x)\big) - xh(a) - (1 - x)h(c)$$

where $h : [0, 1] \to \mathbb{R}$ is the binary entropy function[1]

$$h(x) = -x \log x - (1 - x) \log (1 - x) \tag{1}$$

and $f : [0, 1] \to [0, 1] \subseteq \mathbb{R}$ is $f(x) = (a - c)x + c$. Of course, we can also let a and c vary and view $I(x)$ as a function of three variables $I_x(a, c)$. For fixed a and c, $C = C(a, c) = \max_x I(x) = \max_x I_x(a, c)$, where C is the capacity [22] as a continuous [19, Sec 2.4], [14], [15] function of a and c.

It can be shown (see Silverman [24, Eq. 5], Ash [2, Eq. 3.3.5], or [13], [14]) that capacity $C : I^2 \to [0, 1]$, where $I^2 = [0, 1] \times [0, 1]$, as a function of a and c is

$$\begin{aligned} \mathrm{C}(a, c) &= \tfrac{\bar{a}h(c) - \bar{c}h(a)}{a - c} + \log\left(1 + 2^{\frac{h(a) - h(c)}{a - c}}\right) \\ &= \log\left(2^{\frac{\bar{a}h(c) - \bar{c}h(a)}{a - c}} + 2^{\frac{ch(a) - ah(c)}{a - c}}\right), \end{aligned} \tag{2}$$

where $\mathrm{C}(a, a) := 0$. One also has the symmetries

$$C(a, c) = C(c, a) = C(1 - a, 1 - c) = C(1 - c, 1 - a)\ . \tag{3}$$

$C(a, c)$ is defined and continuous for all $(a, c) \in [0, 1] \times [0, 1]$, is zero for the values (a, a), and is 1 if and only if $(a, c) = (1, 0)$, or $(a, c) = (0, 1)$, see Figure 2.

We let x_0 denote the value of x that maximizes mutual information for a given (a, c), and define $\beta(a, c)$ by

$$\beta(a, c) = \frac{h_e(a) - h_e(c)}{a - c}, \tag{4}$$

where $h_e(x) := -x \ln x - (1 - x) \ln (1 - x) = \ln(2) \cdot h(x)$. The expression[2] for x_0 is given in [2], [24] as:

$$x_0 = \frac{1}{a - c}\left(\frac{1}{1 + e^{\beta}} - c\right),\quad 0 \le a, c \le 1, a \ne c\ . \tag{5}$$

Note that x_0 cannot be continuously extended to the closed unit square $[0, 1] \times [0, 1]$. This is due to the fact that (as first noted by Silverman [24], see also [6], and proved rigorously for more general timing channels in [13])

$$\lim_{a \to 0^+} x_0(a, 0) = 1/e, \quad \lim_{c \to 0^+} x_0(0, c) = 1 - 1/e \tag{6}$$

$$\lim_{a \to 1^-} x_0(a, 1) = 1 - 1/e, \quad \lim_{c \to 1^-} x_0(1, c) = 1/e\ . \tag{7}$$

The above shows that we cannot extend x_0 to be a continuous function including the corner points $(0, 0)$ and $(1, 1)$. Stated, without proof, in [24] is the Silverman conjecture

$$1/e \le x_0 \le 1 - 1/e\ . \tag{8}$$

This was not proved until 1989 by Rumsey, as demonstrated in the dissertation of Majani [11], [12] (See also [23]).

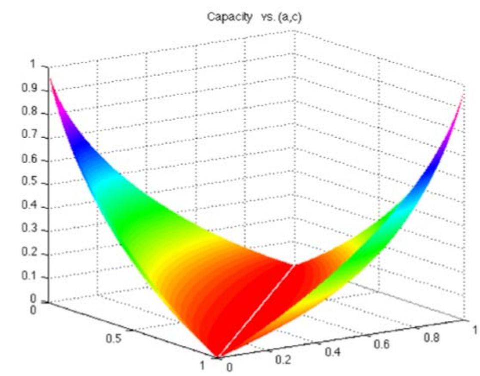

Fig. 2. Capacity as a function of (a, c).

I-B. *Stochastic Resonance in a (2,2) Channel*

Now we discuss how adding noise can counterintuitively increase channel capacity. Note that one must be careful with the definition of noise. The motivation for this comes from the physical world [20].

In terms of Shannon's information theory, a (2,2) channel is noiseless [2] if M is the two by two identity matrix, or its flip (rows interchanged). However, we are now faced with the dilemma of what it means for one channel to be more noisy than another? Do we compare the capacities of the two channels, do we compare the terms a and c, or do we devise some metric to see how "far away" the channel matrix is from the identity matrix, or its flip? No matter what definition of noise we use, it seems counterintuitive for noise to increase capacity. We feel that the use of the term "noise" in stochastic resonance applied to information theory is a misnomer and we would be better served with the term "disturbance noise," or simply "disturbance," denoted by $\mathfrak{N}_{\mathfrak{D}}$. The idea of disturbance agrees with the communication system model as shown in [1, Fig. 3.1]. It is the addition of $\mathfrak{N}_{\mathfrak{D}}$ that can increase channel capacity.

The input random variable X represents the physical transmission of a signal $\mathfrak{X}$. In our situation $\mathfrak{X}$ takes on two distinct values, level 0 or level 1, represented in terms of the random variable as ι_1 or ι_2, respectively. The disturbance is added to

[1]Note, "log" is base two, whereas as "ln" is the natural logarithm. The binary entropy function is defined for $x \in [0, 1]$, with $h(0) = h(1) := 0$.

[2]It is worth noting that $e^{\beta(a,c)} = 2^{\frac{h(a)-h(c)}{a-c}}$.

the transmission signal resulting in the received signal $\mathfrak{R}$:

$$\mathfrak{R} = \mathfrak{X} + \mathfrak{N}_{\mathfrak{D}} \; . \tag{9}$$

That is, each time an input symbol is sent, an independent draw from $\mathfrak{N}_{\mathfrak{D}}$ is added to it. Thus, the disturbance may increase/decrease, or leave unchanged, the transmitted signal level. The value (level) of $\mathfrak{R}$ can be any real number. However, as in [3], the concept of a threshold is developed based on the physical model. The threshold interprets $\mathfrak{R}$, resulting in an output random variable Y in a discrete Shannon channel. We do not analyze the physical reality of such output thresholding [10], [3]; rather we concentrate on the information theoretic analysis of such a scenario (information theoretical analysis of spike trains/action potentials is a rich area of research, *e.g.* [18]). We optimistically and humbly suggest that a judicious addition of noise into a synapse signal might assist in neurological disease healing (for background see [4]).

We have a threshold[3] θ such that

$$Y = \begin{cases} 0, & \text{if } \mathfrak{R} \le \theta; \\ 1, & \text{if } \mathfrak{R} > \theta. \end{cases} \tag{10}$$

Following [3], we let $\mathfrak{N}_{\mathfrak{D}}$ being a zero mean ($\mu = 0, \sigma > 0$) Gaussian distribution, with probability density function $f_\sigma(t) = \frac{1}{\sqrt{2\pi}\cdot\sigma}\exp(\frac{-t^2}{2\sigma^2})$, and cumulative distribution function $F_\sigma(x) = \int_{-\infty}^{x} f_\sigma(t)dt$. We denote the cumulative distribution function of the standard normal distribution ($\mu = 0,\ \sigma = 1$) as

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}}\, dt \; . \tag{11}$$

Change of variables shows that $F_\sigma(x) = \Phi(x/\sigma)$. Therefore, we easily have [3] that the channel matrix is of the form[4]

$$\begin{aligned} M = \begin{pmatrix} a & 1-a \\ c & 1-c \end{pmatrix} &= \begin{pmatrix} F_\sigma(\theta) & 1 - F_\sigma(\theta) \\ F_\sigma(\theta-1) & 1 - F_\sigma(\theta-1) \end{pmatrix} \\ &= \begin{pmatrix} F_\sigma(\theta) & \overline{F_\sigma(\theta)} \\ F_\sigma(\theta-1) & \overline{F_\sigma(\theta-1)} \end{pmatrix}, \end{aligned} \tag{12}$$

and the channel capacity is therefore

$$C\left(F_\sigma(\theta), F_\sigma(\theta-1)\right) = \tag{13}$$

$$\log\left(2^{\frac{\overline{F_\sigma(\theta)}\cdot h(F_\sigma(\theta-1)) - \overline{F_\sigma(\theta-1)}\cdot h(F_\sigma(\theta))}{F_\sigma(\theta)-F_\sigma(\theta-1)}} + 2^{\frac{F_\sigma(\theta-1)\cdot h(F_\sigma(\theta)) - F_\sigma(\theta)\cdot h(F_\sigma(\theta-1))}{F_\sigma(\theta)-F_\sigma(\theta-1}}\right) \; .$$

When we wish to emphasize that we are analyzing $C(F_\sigma(\theta), F_\sigma(\theta-1))$ with σ fixed, we slightly abuse notation and denote the capacity as $C_\sigma(\theta)$.

Note that we also have the trivial identity of zero mean Gaussian distributions

$$F_\sigma(-\theta) = 1 - F_\sigma(\theta) \; . \tag{14}$$

[3]Of course, we could simplify the exposition by using values centered about 0. However, we have used the present formulation to mimic [3].

[4]Where $\bar{x} = 1 - x$.

Intuitively, raising the threshold from 1 to $1 + \Delta$ should, from the point of view of information theory, have the same effect as dropping it from 0 to $-\Delta$. One swaps the input symbols and also the output symbols. Some more thought shows that the symmetry is actually between 0 and 1; that is the channel capacity should be symmetric about 1/2. We prove this (using Eqs. 3 & 14) mathematically as well; that is, we will show that $C_\sigma(.5 + d) = C_\sigma(.5 - d)$.

Theorem 1.1: $C_\sigma(.5 + d) = C_\sigma(.5 - d)$.

Proof: $C_\sigma(\theta) = C\left(F_\sigma(\theta), F_\sigma(\theta-1)\right) = C\left(F_\sigma(\theta-1), F_\sigma(\theta)\right) = C\left(1 - F_\sigma(1-\theta), 1 - F_\sigma(-\theta)\right) = C\left(F_\sigma(1-\theta), F_\sigma(-\theta)\right) = C_\sigma(1-\theta)$, so $C_\sigma(.5 + d) = C_\sigma\left(1 - (.5 + d)\right) = C_\sigma(.5 - d)$ ■

Nota Bene: Therefore, without loss of generality, we will always assume that $\theta \ge .5$, since we can always obtain the values for $\theta < .5$ by symmetry.

In [3] the stochastic resonance effect was demonstrated for $\theta > 1$, and no stochastic resonance was demonstrated for $.8 < \theta \le 1$. This was done via discussion and plots. Specifically, it was discussed and demonstrated that for $.8 < \theta \le 1$, the capacity monotonically decreases from 1 to 0, as a function of $\sigma > 0$. This is because if the threshold $\theta < 1$ then we have a noiseless and useless channel if there is no added disturbance. Note, that in the limit as $\sigma \to 0^+$ that $\mathfrak{N}_{\mathfrak{D}} \to \delta(t)$, the Dirac impulse function at 0. We note that [3] did not discuss the specifics of $\theta < 1$, the lessons learned in [3] only hold if $.5 \le \theta < 1$. (See also the "forbidden interval" theorem [8, Sec. 6.2].)

If there is no added disturbance, when the threshold $\theta > 1$, the threshold will never be met and there is no information flowing; hence, the capacity is 0. However, as soon as there is any non-trivial added disturbance (that is $\sigma > 0$), there will be non-zero mutual information, and hence non-zero capacity. Graphically, it is shown in [3], that the capacity starts at 0 for $\sigma = 0$ and then increases as σ increases. However, when the added disturbance grows too large, the capacity starts to decrease again and approach 0. This makes sense because $F_\sigma(\theta) = \Phi(\frac{\theta}{\sigma})$, and $F_\sigma(\theta-1) = \Phi(\frac{\theta}{\sigma} - \frac{1}{\sigma})$. So for large σ, $F_\sigma(\theta-1) \to F_\sigma(\theta)^-$, and hence $C \to 0$.

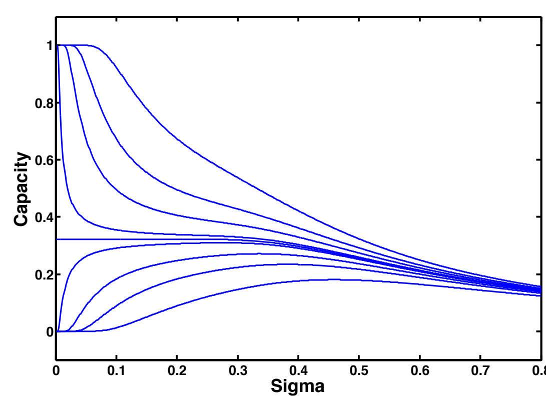

Fig. 3. Capacity, for $\theta \in \{.8, .9, .95, .99, 1, 1.01, 1.05, 1.1, 1.2\}$, as a function of σ.

We reproduce in Figure 3 the same plots that are given in [3]. In these plots $\sigma \in (0, .8]$ and $\theta \in \{.8, .9, .95, .99, 1, 1.01, 1.05, 1.1, 1.2\}$. The θ values plot monotonically from the top (.8) to the bottom (1.2). As discussed above, stochastic resonance is not exhibited until $\theta > 1$. For $\theta \leq 1$ the capacity decreases monotonically from 1, and for $\theta > 1$ the capacity increases from 0 to a maximum value and then decreases.

Definition 1.2: For a given θ, a σ value that locally maximizes capacity is called a *local harmonic*, if it is a global maximum, we simply call it the *harmonic*, and denote it as σ^h.

Judging from Figure 4, and our discussions, there appear to be only harmonics, and that for $\theta > 1$. But, how do we prove this? The obvious tack would be, for fixed θ, to express capacity as a function of σ, written as $C_\theta(\sigma)$, and then to analyze $\frac{d}{d\sigma}C_\theta(\sigma)$. However, that does not seem to be readily tractable, as can be seen by examining the closed form for capacity as given in Eqs. (2,13). Instead, we propose using approximations to the capacity which we discuss below.

II. Capacity Bounds and Approximations

So far, all of our discussions of capacity have been performed by using the closed form expressions Eqs. (2,13). However, the closed form for capacity is very difficult to deal with because of its non-linear behavior. Also, the closed form does not generalize (unless the channel matrix is square and invertible) to higher dimensional channels, which is a topic for future research (see also [3]). It was because of the mathematical awkwardness of the capacity expression that approximations and bounds were derived in [16], [17].

A $(2, n)$ channel is a discrete memoryless Shannon communication channel with two inputs and n outputs. Its (stochastic) channel matrix is $M = \begin{pmatrix} a_1 & \dots & a_{n-1} & \bar{a} \\ c_1 & \dots & c_{n-1} & \bar{c} \end{pmatrix}$ where $\bar{a} = 1 - (a_1 + \cdots + a_{n-1})$ and $\bar{c} = 1 - (c_1 + \cdots + c_{n-1})$.

We use the following approximations to capacity; we call the lower bound the *Pinsker bound*, and the upper bound the *Helgert bound*. These bounds, and their use as capacity approximations, and the following theorem, are discussed in detail in [16], [17].

Theorem 2.1: For a $(2, n)$ channel, the capacity has Pinsker (lower) bound and Helgert (upper) bound as given below

$$\begin{aligned} \frac{1}{8\ln(2)} \left[\left\{ \sum_{i=1}^{n-1} |a_i - b_i| \right\} + |\bar{a} - \bar{b}| \right]^2 \quad &\leq \quad C \\ \leq \quad \max \Big(\sum_{i=1}^{n-1} a_i, \sum_{i=1}^{n-1} b_i \Big) - &\sum_{i=1}^{n-1} \min(a_i, b_i) \quad . \quad (15) \end{aligned}$$

For a $(2, 2)$ channel the above reduces to:

$$\frac{(a-b)^2}{2\ln(2)} \leq C \leq |a - b| \ . \qquad (16)$$

We note that the determinant of the channel matrix is $a - b$ [14], and $a - b$ is also an eigenvalue of the channel matrix [5].

III. Approximation Harmonics for the $(2, 2)$ Channel

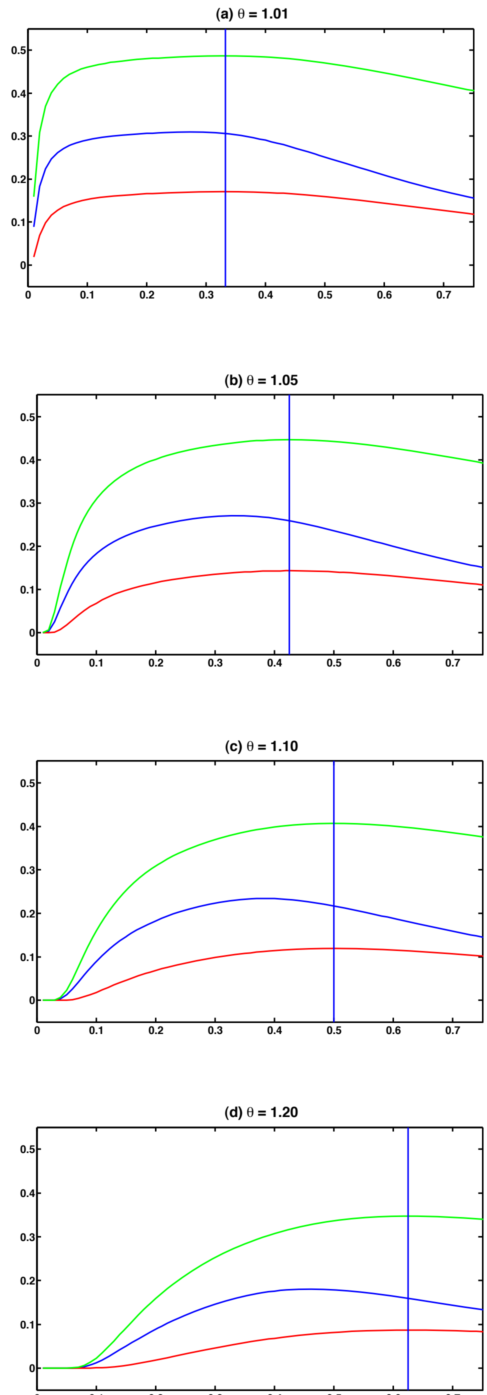

Fig. 4. Approximations to σ^h: where vertical lines intersect top and bottom plots.

From Eq. (16), and the fact $F_\sigma(\theta) \geq F_\sigma(\theta-1)$ when $\sigma > 0$ (Note that $F_\sigma(\theta) \geq F_\sigma(\theta-1)$ for $\sigma \geq 0$), we have

$$\frac{\left(F_\sigma(\theta) - F_\sigma(\theta-1)\right)^2}{2\ln(2)} \leq \quad C_\sigma(\theta) \quad \leq \quad F_\sigma(\theta) - F_\sigma(\theta-1)\ .$$

Note that these bounds share the symmetry about $\theta = .5$ with the capacity. This is because $F_\sigma(.5+d) - F_\sigma(.5+d \quad -1) = F_\sigma(.5+d) - F_\sigma(-.5+d) = 1 - F_\sigma(-.5-d) - (1-F_\sigma(.5-d)) = F_\sigma(.5-d) - F_\sigma(-.5-d) = F_\sigma(.5-d) - F_\sigma(.5-d \quad -1)$. Therefore, we can continue with our blanket statement that $\theta \geq .5$.

In Figure 4 we plot the Pinsker bound in red, the capacity in blue, and the Helgert bound in green. The bounds also only exhibit stochastic resonance for $\theta > 1$. Since the bounds are much more tractable than the closed form for capacity, we instead find the harmonics of the bounds.

From above we have that the (lower) Pinsker bound is

$$\mathfrak{P} = \frac{\left(F_\sigma(\theta) - F_\sigma(\theta-1)\right)^2}{2\ln(2)}$$

and the (upper) Helgert bound is $\mathfrak{H} = F_\sigma(\theta) - F_\sigma(\theta-1)$. We are interested in the above critical points of the bounds.

So we now consider $\frac{d}{d\sigma}\mathfrak{P}$ and $\frac{d}{d\sigma}\mathfrak{H}$ (we take θ as fixed and view the bounds as only functions of $\sigma > 0$). The critical points of $\frac{d}{d\sigma}\mathfrak{P}$ coincide with the critical points of $\frac{d}{d\sigma}\mathfrak{H}$, so we only concern ourselves with the later.

$$\begin{aligned}
\frac{d}{d\sigma}\mathfrak{H} &= \frac{d}{d\sigma}\left(\frac{1}{\sqrt{2\pi}\sigma}\int_{\theta-1}^{\theta} e^{\frac{-t^2}{2\sigma^2}}\,dt\right) \\
&= \frac{-1}{\sqrt{2\pi}\sigma^2}\left(\int_{\theta-1}^{\theta} e^{\frac{-t^2}{2\sigma^2}}\,dt + \int_{\theta-1}^{\theta} t\cdot\frac{-t}{\sigma^2}e^{\frac{-t^2}{2\sigma^2}}\,dt\right) \\
&= \frac{-1}{\sqrt{2\pi}\sigma^2}\left(\int_{\theta-1}^{\theta} e^{\frac{-t^2}{2\sigma^2}}\,dt + \int_{\theta-1}^{\theta} t\cdot\frac{d}{dt}\left(e^{\frac{-t^2}{2\sigma^2}}\right)dt\right) \\
&= \frac{-1}{\sqrt{2\pi}\sigma^2}\left(\int_{\theta-1}^{\theta} e^{\frac{-t^2}{2\sigma^2}}\,dt + te^{\frac{-t^2}{2\sigma^2}}\Big|_{\theta-1}^{\theta} - \int_{\theta-1}^{\theta} e^{\frac{-t^2}{2\sigma^2}}dt\right) \\
&= \frac{-1}{\sqrt{2\pi}\sigma^2}\left(\theta e^{\frac{-\theta^2}{2\sigma^2}} - (\theta-1)e^{\frac{-(\theta-1)^2}{2\sigma^2}}\right).
\end{aligned}$$

Setting this equal to zero we arrive at

$$\frac{\theta}{\theta-1} = e^{\frac{2\theta-1}{2\sigma^2}}. \tag{17}$$

We restrict θ such that $\theta > 1$, and are use $\sigma > 0$

$$\sigma = \sqrt{\left(\theta - \frac{1}{2}\right) \Big/ \left(\ln\left(\frac{\theta}{\theta-1}\right)\right)}\ . \tag{18}$$

This is an approximation, which we denote as $\sigma^{\widetilde{h}}$ for a given θ, to the optimal noise power.

We see in Figure 5, the limiting behavior of $\sigma^{\widetilde{h}}$. This limiting behavior is trivial to show from the above equation, but we wish to note that it coincides with our physical intuition.

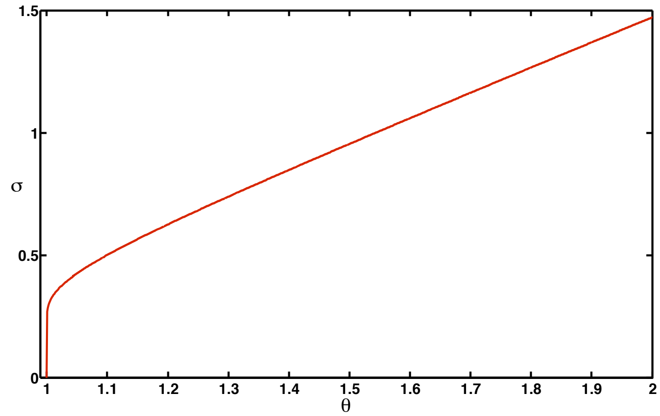

Fig. 5. Optimal noise power approximation $\sigma^{\widetilde{h}}$ as a function of θ.

IV. Conclusion and Future Work

We have given physical motivation for our information theoretic analysis. We have discussed closed form solutions in terms of channel matrix elements and have shown how capacity bounds can readily show how matrix terms affect capacity. An immediate lesson learned from this paper is that adding noise with a fixed mean may be counter-productive if one wishes to negate transmission, and we should rather add noise with various means or other differences in moments. This will be a topic for future research along with a further analysis via algebraic information theory. In addition, in future work, we will use simulations to see how capacity might actually be achieved in both the theoretical maximums and out approximate maximums

We also wish to keep in mind how stochastic resonance may affect jamming, and to consider, from how the ideas from stochastic resonance, may be used to improve signal transmission in defective neural synapses. In addition in future work, we will use simulations to investigate jamming and to see how capacity might actually be achieved in both the theoretical maximums and our approximate maximums.

Acknowledgments

We thank François Chapeau-Blondeau for his helpful email correspondence. The second author acknowledges a 2011 ONR-ASEE Fellowship and DOD Grant W911NF-11-1-0144. We are grateful to Moshe Ben-Maimon for his helpful explanations. We thank R. Shulamite and the reviewers for their helpful comments.

References

[1] John M. Wozencraft amd Irwin Mark Jacobs. *Principles of Communication Engineering*. Wiley, 1965.

[2] Robert B. Ash. *Information Theory*. Dover, 1990.

[3] François Chapeau-Blondeau. Noise-enhanced capacity via stochastic resonance in an asymmetric binary channel. *Physical Review E*, 55(2):2016 – 2019, 1997.

[4] Edward C. Cooper and Lily Y. Jan. M-channels neurological diseases, neuromodulation, and drug development. *Archives of Neurology*, 60:496–500, April 2003.

[5] Paul Cotae, Ira S. Moskowitz, and Myong H. Kang. Eigenvalue characterization of the capacity of discrete memoryless channels with invertible channel matrices. In *Proc. 44th Annual Conf. on Information Sciences and Systems, CISS 2010*, Princeton, NJ, USA, March 2010.

[6] Solomon W. Golomb. The limiting behavior of the Z-channel. *IEEE Transactions on Information Theory*, 66(3):372, May 1980.

[7] J.J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the USA*, 79:2554–2558, April 1982.

[8] Bart Kosko. *Noise*. Viking, 2006.

[9] Bart Kosko and Sanya Mitaim. Stochastic resonance in noisy threshold neurons. *Neural Networks*, 16:755–761, 2003.

[10] Jacon E. Levin and John P. Miller. Broadband neural encoding in the cricket cereal sensory system enhanced by stochastic resonance. *Nature*, 380:165–168, March 1996.

[11] E. E. Majani. *A Model for the Study of Very Noisy Channels & Applications, 1988*. PhD thesis, Cal Tech.

[12] E.E. Majani and H. Rumsey. Two results on binary-input discrete memoryless channels. In *Proceedings Int. Symp. on Information Theory*, page 104, 1991.

[13] Keye Martin and Ira S. Moskowitz. Noisy timing channels with binary inputs and outputs. In Camenisch et al., editor, *Information Hiding, 8th International Workshop*, pages 124–144, Alexandria, VA, USA, July 2006. Springer-Verlag, LNCS 4437.

[14] Keye Martin, Ira S. Moskowitz, and Gerard Allwein. Algebraic information theory for binary channels. In *Proceedings MFPS XXII*, Genova, Italy, May 2006. ENTCS, V. 158, pp. 289–306.

[15] Keye Martin, Ira S. Moskowitz, and Gerard Allwein. Algebraic information theory for binary channels. *Theoretical Computer Science*, 411:1918–1927, April 2010.

[16] Ira S. Moskowitz. An approximation of the capacity of a simple channel. In *Proc. 43rd Annual Conf. on Information Sciences and Systems, CISS 2009*, pages 164–169, Baltimore, MD, USA, March 2009.

[17] Ira S. Moskowitz. Approximations for the capacity of binary input discrete memoryless channels. In *Proc. 44th Annual Conf. on Information Sciences and Systems, CISS 2010*, Princeton, NJ, USA, March 2010.

[18] Ira S. Moskowitz and Richard E. Newman. Timing channels, anonymity, mixes, and spikes. In *ACST'06*, pages 252–256. IASTED, 2006.

[19] Ira S. Moskowitz, Richard E. Newman, Daniel P. Crepeau, and Allen R. Miller. Covert channels and anonymizying networks. In *Proc. ACM Workshop on Privacy in Electronic Society (WPES)*, pages 79–88. ACM, October 2003.

[20] Frank Moss, David Pierson, and David O'Gorman. Stochastic resonance: Tutorial and update. *Int. J. of Bifurcation and Chaos*, 4(6):1383–1397, 1994.

[21] R. Schroer. Electronic warfare. *Aerospace and Electronic Systems Magazine, IEEE*, 18(7):49 – 54, July 2003.

[22] Claude E. Shannon. A mathematical theory of communication. *Bell Systems Technical Journal*, 27:379–423,623–656, 1948.

[23] Nadav Shulman and Meir Feder. The uniform distribution as a universal prior. *IEEE Transactions on Information Theory*, 50(6):1356–1362, June 2004.

[24] Richard A. Silverman. On binary channels and their cascades. *IRE Transactions on Information Theory*, 1(3):19–27, May 1955.